

- 14 -
CLAIMS

1. A method for handling failover of a data management application for a shared disk file system in a distributed computing environment having a cluster of loosely coupled nodes which provide services, comprising the steps of:

defining certain nodes of the cluster as failover candidate nodes;

storing configuration information for all the failover candidate nodes;

distributing message information including but not limited to failure information of at least one failover candidate node amongst the failover candidate nodes;

analyzing the distributed message information and the stored configuration information in order to determine whether to take over the service of a failure node by a failover candidate node or not;

updating the configuration information in case of at least one failover candidate node taking over the service of a failure node.
2. Method according to claim 1, wherein a failover request is carried on to at least a second failover candidate if only a subset of the filesystem is taken over from the failure node by a first failover candidate.
3. Method according to claim 1 or 2, wherein the configuration information is stored in a central data storage arranged within the cluster.
4. Method according to any of claims 1 to 3, wherein the distributed message information includes a failure report of at least one node.

5. Method according to any of the preceding claims, wherein the failover candidate nodes calculate a priority key related to the workload of each of the failover candidate nodes which is distributed as part of the distributed message information.
6. Method according to claim 5, wherein the failover candidate nodes receiving the priority key compare the received priority key with their own priority key whereby the best priority key wins the right to take over the service.
7. Method according to any of the preceding claims, wherein the updating of the configuration information is handled by means of a locking mechanism.
8. An article of manufacture comprising a computer usable medium having computer readable program code means embodied therein for causing handling failover of a data management application for a shared disk file system in a distributed computing environment having a cluster of loosely coupled nodes which provide services, the computer readable program code means in the article of manufacture comprising computer readable program code means for causing a computer to effect:

defining certain nodes of the cluster as failover candidate nodes;

storing configuration information for all the failover candidate nodes;

distributing message information including but not limited to failure information of at least one failover candidate node amongst the failover candidate nodes;

analyzing the distributed message information and the stored configuration information in order to determine whether to take over the service of a failure node by a failover candidate node or not;

updating the configuration information in case of at least one failover candidate node taking over the service of a failure node.

9. A system for handling failover of a data management application for a shared disk file system in a distributed computing environment having a cluster of loosely coupled nodes which provide services, comprising

data storage means for storing configuration information for failover candidate nodes;

communication interface means for distributing message information between the failover candidate nodes;

means for analyzing the message information and the configuration information in order to determine whether to take over the service of a failure node by a failover candidate node or not;

means for updating the configuration information in case of at least one failover candidate node taking over the service of a failure node.

10. System according to claim 9, further comprising means for cascading the failover handling whereby a failover request is carried on to at least a second failover candidate if only a subset of the filesystem is taken over from the failure node by a first failover candidate.
11. System according to claim 9 or 10, where the data storage means is a central data storage arranged within the cluster.
12. System according to any of claims 9 to 11, where the means for updating the configuration information are located at the failover candidate node taking over a service of a failure node.